DOCUMENT RESUME

ED 452 269 TM 032 557

AUTHOR Gorney, Barbara; Maury, Marcia

TITLE Relationship of Scores and Times of Test Administration via Computer.

PUB DATE 2001-04-00

NOTE 8p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Computer Assisted Testing; High Stakes Tests; Higher Education; *Licensing Examinations (Professions); Medical Education; *Medical Students; *Scores; *Test Results; *Time

ABSTRACT

The purpose of this study was to determine if a relationship exists between scores and the times that medical students choose to take a computer-administered test. The results indicate that students who choose to take a test later within a given time period tend to perform less well than students who take the test earlier. Although the magnitude of the relationships was slight, a negative association was found for both high-stakes national examinations and weekly quizzes associated with a medical school course. Second-year students are required to take the U.S. Medical Licensing Examination (USMLE) Step 1 in March or April, and third-year students are required to take Step 2 between April and July. Students, who are required to attain a national passing score on both examinations in order to progress, may schedule their examinations from several weeks to several months in advance. Data from the Step 1 examination were studied for 2 cohorts of 111 and 102 students. Although the magnitude of the associations found was slight, the negative direction persisted across cohorts and type of test, whether high-stakes national tests or weekly quizzes for second-year students. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Relationship of Scores and Times of Test Administration via Computer

Barbara Gorney
Marcia Maury

ED 452 269

TM032557

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
B. Gorney
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

This paper is prepared for the:
Annual Meeting of the American Educational Research Association in Seattle, WA
April 2001

BEST COPY AVAILABLE

Relationship of Scores and Times of Test Administration via Computer

Barbara Gorney and Marcia Maury

Presented at the American Educational Research Association Annual Meeting, Seattle, April, 2001.

Abstract

The purpose of this study was to determine if a relationship exists between scores and the times that medical students chose to take a computer-administered test. The results obtained indicated that students who chose to take a test later within a given time period tend to perform less well than students who took the test earlier. Although the magnitude of the relationships was slight, a negative association was found for both high-stakes national examinations and weekly quizzes associated with a medical school course.

Introduction

Administration of student assessments via computer is becoming more commonplace, not only among national testing organizations but also in a variety of educational and corporate settings. One of the positive aspects of computer-based testing reported by students is being able to schedule a test when they believe they are ready (Bugbee & Bernt, 1990). Computer testing software typically has the capability to record information about the testing session itself, for example, the length of time taken to complete a test and the date and time of testing session. Information pertaining to a student's testing session may provide educators with information that may be used to improve student learning and performance.

Medical students are typically required to take the USMLE Step 1 and Step 2 licensure examinations at the end of the second and third years of medical school, respectively. Beginning in 1999, these examinations moved from paper-and-pencil administrations on set dates to being computer-administered at a time and place individually scheduled by the students. At Wake Forest University School of Medicine,

anecdotal evidence was put forth that students who scheduled their exams late in the testing period tended to perform poorly.

The Wake Forest University School of Medicine issues a laptop computer to all incoming medical students. Several courses administer summative weekly quizzes and major examinations via LAN-based testing software. Students use their laptops to take these quizzes and exams at a time of their choosing within established time parameters, typically 5 to 7 days.

A study was undertaken to determine if a relationship existed between scores and the times that medical students chose to take an assessment within a defined period of time. Both the national, standardized licensure examinations and the course-level weekly quizzes were examined. Neither type of test is computer-adaptive.

A search of the education and medical literature did not reveal any studies relating student performance to when students sat for an examination. Studies were found relating performance to the amount of time spent on tests (e.g., Kennedy, 1994), time of day (Callan, 1995), and time-related testing behaviors (Lin, 1986).

Methodology

The subjects were students at one medical school. The academic schedule for a cohort or class of medical students is different from that of a cohort of college students because, generally, medical students all attend the same classes at the same times. The confounding effect of different class schedules was essentially controlled in this study.

Second-year medical students took weekly computer-administered quizzes in two courses. The quiz scores contributed substantially toward the course grade. Quizzes were placed on a central server location each Friday and remained available to students

for the next five days. Students used their individual laptops to take the quiz but had to be on-campus to access the server.

The session-related information recorded by the testing software includes the beginning date and time, the ending date and time, and total time taken to complete the test for each student along with the score. Date and time of test-taking were rank ordered so that the first person to take the quiz received a rank of 1. The date/time rank ordering was correlated with the quiz scores.

Second-year medical students are required to take the US Medical Licensing Examination (USMLE) Step 1 in March or April. Third-year students are required to take the USMLE Step 2 between April and July. Students at the medical school must attain a national passing score on both exams in order to progress. Students may schedule their examinations from several weeks to several months in advance.

Included in the score information provided to medical schools by USMLE is the date on which the student took the exam. Data for two student cohorts on the Step 1 and one cohort on the Step 2 were examined to determine if a relationship existed between testing dates and scores. Test dates were rank ordered with the first date receiving a rank of 1. There were a substantial number of test dates with tied ranks.

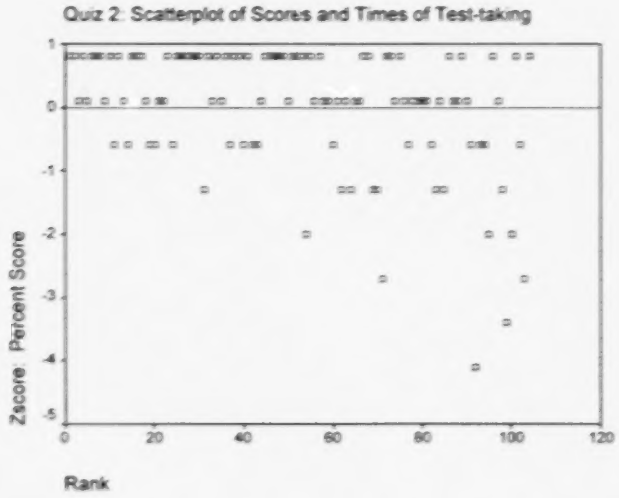
Results

Students took a total of six quizzes during a 9-week period; quiz 1-4 and quiz 5-6 were on different topics (neurophysiology and hematology, respectively). Each relationships observed between quiz scores and times of test was slightly to moderately negative. As shown in Table 1, Pearson correlation coefficients ranged from –0.40 to –0.21.

Table 1. Correlation of quiz score and rank-ordered time of testing for six course-related quizzes by one cohort of students.

	Pearson r	n
Quiz 1	-.32	107
Quiz 2	-.40	104
Quiz 3	-.34	107
Quiz 4	-.23	107
Quiz 5	-.21	107
Quiz 6	-.23	106

Examination of the scatterplots of these data indicate that the negative relationships are due to a greater proportion of students who did poorly taking the quizzes later in the time period. It appears that students who do well are just as likely to take quizzes early as to take them late. The scatterplot of Quiz 2 which provides the most dramatic example is shown below.



Data from the USMLE Step 1 were examined for two cohorts of second-year medical students. The relationship between test scores and test dates was slightly negative for both classes. Pearson correlation coefficients were -.28 ($n = 111$, $p < .01$) and -.25 ($n = 102$, $p < .05$). The relationship between Step 2 scores and test dates for one

cohort of third year medical students was also slightly negative but not statistically significant ($r = -.17$, $n = 103$, $p > .05$).

Data were also examined for one course-end, flexible schedule computer-administered test given during a 7-day period when the students in the cohort did not have the same academic schedules. The relationship between test scores and times of test-taking was still negative but very weak (Pearson's $r = -.08$, $n = 106$). This result may support the notion that differing academic schedules is a confounding factor.

Discussion/Conclusions

The results obtained in this study showed a negative association between performance and the time students chose to take an exam or quiz during a finite testing period. Although the magnitude of the associations was slight, the negative direction persisted across student cohorts and type of test, specifically, the high stakes national USMLE Step 1 and Step 2 and weekly summative quizzes associated with second year medical school courses. Examination of the data scatterplots indicated that students who scored lower tended to take tests later during the prespecified time period.

Faculty may use the results found in this study for weekly quizzes to identify struggling students early in a course and then to help students redirect their study habits, before the major examinations. Students who are struggling, in particular, would likely benefit from studying the material presented on a daily basis. Review of the week's material and completion of the quiz prior to the start of a new week should help students from becoming overwhelmed by material to be learned.

The small, negative correlations obtained for two years of Step 1 scores were virtually identical. Many medical schools require passing scores on the Step 1 and Step 2

national licensure examinations prior to graduation. Administrators regularly receive information regarding their students' Step exams scheduling, including when application was made and the scheduled test date. Review of this type of scheduling information could allow medical schools to offer additional preparatory support to students who plan to take these exams at the end of the testing period.

References

Bugbee, A.C., Jr., and Bernt, F.M. (1990). Testing by Computer: Findings in Six Years of Use 1982-1988. *Journal of Research on Computing in Education*, pp. 87-100.

Callan, R.J. (1995). Early Morning Challenge: The Potential Effects of Chronobiology on Taking the Scholastic Aptitude Test. *Clearing House*, pp. 174-76.

Kennedy, R. (November, 1994). A Study of the Relationship between Scores and Time on Tests. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Nashville.

Lin, M. (April, 1986). The Impact of Time Limits on Test Behaviors. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.